



PATENT ABSTRACTS OF JAPAN

(11) Publication number: 2003208199 A

(43) Date of publication of application: 25.07.03

(51) Int. Cl. G10L 19/00
H03M 7/38

(21) Application number: 2002005562

(22) Date of filing: 15.01.02

(71) Applicant: HITACHI ULSI SYSTEMS CO LTD

(72) Inventor: MORI TAKEO
SEO MICHIIKO
TANAKA MAKOTO

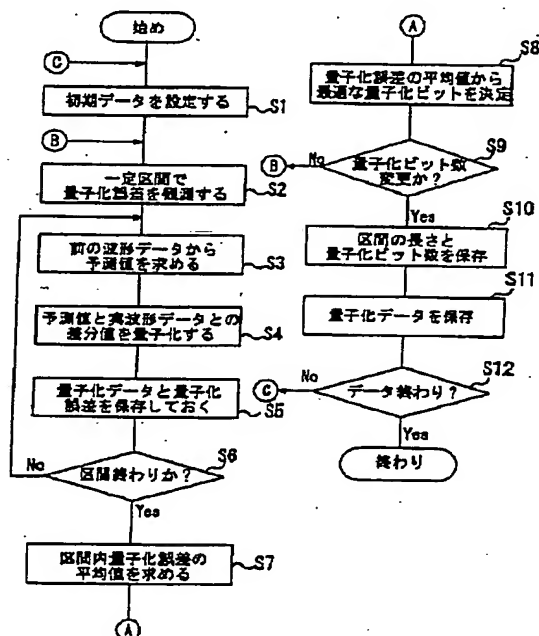
(54) VOICE COMPRESSING METHOD AND
EXTENDING METHOD

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a voice compressing method and extending method by which the deterioration of a sound quality is reduced and an entire data compression rate can be improved, and which can be carried out by inexpensive hardware constitution.

SOLUTION: The voice compressing method samples a voice signal, thereafter predicts a sample value at each time point based on sample values before then and quantizes and encodes a difference between this predicted value and an actual sample value to generate voice compression data. A series of voice signals are divided by each small section and respective small sections are quantized so that the average of the encoding errors of the small sections may be within a prescribed range. When the average of the encoding errors exceeds the prescribed range, the number of quantizing bits is varied.

COPYRIGHT: (C)2003,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-208199

(P2003-208199A)

(43) 公開日 平成15年7月25日 (2003.7.25)

(51) Int.Cl.⁷

識別記号

F I

テ-マ-ト* (参考)

G 1 0 L 19/00

H 0 3 M 7/38

5 D 0 4 5

H 0 3 M 7/38

G 1 0 L 9/18

C 5 J 0 6 4

審査請求 未請求 請求項の数 5 O L (全 7 頁)

(21) 出願番号 特願2002-5562 (P2002-5562)

(22) 出願日 平成14年1月15日 (2002.1.15)

(71) 出願人 000233169

株式会社日立超エル・エス・アイ・システムズ

東京都小平市上水本町5丁目22番1号

(72) 発明者 森 竹雄

東京都小平市上水本町5丁目22番1号 株式会社日立超エル・エス・アイ・システムズ内

(74) 代理人 100085811

弁理士 大日方 富雄

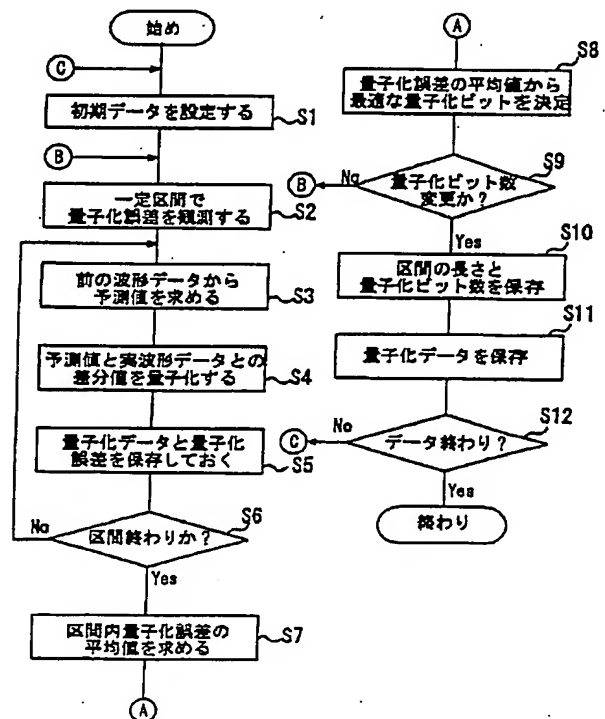
最終頁に続く

(54) 【発明の名称】 音声圧縮方法および伸長方法

(57) 【要約】

【課題】 音質の劣化が少なく且つ全体的なデータ圧縮率を高めることが可能で、且つ廉価なハードウェア構成でも実行することが可能な音声圧縮方法および伸長方法を提供する。

【解決手段】 音声信号を標本化した後、各時点の標本値をそれ以前の標本値に基づき予測するとともに、この予測した値と実際の標本値との差分を量子化および符号化して音声圧縮データを生成する音声圧縮方法において、一続きの音声信号を微小区間ごとに分け、該微小区間の符号化誤差の平均値が所定の範囲に収まるように各微小区間の量子化を行うとともに、符号化誤差の平均値が所定の範囲を越えてしまう場合に量子化ビット数を変動させるように構成する。



【特許請求の範囲】

【請求項 1】 音声信号を標本化し、各時点の標本値をそれ以前の標本値に基づき予測するとともに、この予測した値と実際の標本値との差分を量子化および符号化して音声圧縮データを生成する音声圧縮方法であって、一続きの音声信号を複数に区分けし、各区分毎に異なるビット数で量子化を行うようにしたことを特徴とする音声圧縮方法。

【請求項 2】 一続きの音声信号を微小区分ごとに分け、各微小区分における上記差分の符号化誤差の平均値が所定の範囲に収まるように上記量子化のビット数を変化させることを特徴とする請求項 1 記載の音声圧縮方法。

【請求項 3】 同一の量子化ビット数で連続的に上記差分の量子化が行われている区間の長さ情報と、該区間の量子化ビット数を示す情報とを音声圧縮データに付加することを特徴とする請求項 1 又は 2 に記載の音声圧縮方法。

【請求項 4】 上記微小区分における上記差分の量子化および符号化を前の区間で求めた最適量子化ビット数で行う量子化ステップと、当該微小区分における上記差分の符号化誤差の平均値を求める平均誤差算出ステップと、当該符号化誤差の平均値と当該微小区分で用いられた量子化ビット数とに基づき当該微小区分において符号化誤差の平均値が所定範囲内に収まる最適量子化ビット数を求めるビット数算出ステップとを含み、上記複数の微小区分の各々について時系列順に上記量子化ステップと平均誤差算出ステップとビット数算出ステップとを繰り返すとともに、その間、上記ビット数算出ステップで求めた最適量子化ビット数が変化した場合に、該微小区分を量子化ビットが変化する区分箇所とし、且つ、前の区間において連続的に同一の量子化ビット数で量子化されてきた区間長と当該量子化ビット数の情報をヘッダ情報として音声圧縮データに付加するように構成されていることを特徴とする請求項 2 記載の音声圧縮方法。

【請求項 5】 標本化された音声信号の各時点の標本値が、該標本値とそれ以前の標本値に基づく予測値との差分として表わされるとともに、この差分が量子化および符号化されて保存され、且つ、上記差分の量子化ビット数が複数の区間において異なる値に設定されるとともに、同一の量子化ビット数で連続的に量子化されている区間長と該量子化ビット数の情報とがヘッダ情報として付加されている音声圧縮データを受けて音声信号に伸長する音声伸長方法であって、上記ヘッダ情報から上記区間長と量子化ビット数との情報を読み込む区間情報読込ステップと、当該区間長と量子化ビット数との情報に基づき当該区間長に対応する区間の音声圧縮データを上記量子化ビット数に従って伸長する伸長ステップとを備えていることを特徴とする音声

伸長方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 この発明は、音声の圧縮伸長技術に関し、例えば適応差分パルス符号変調 (ADPCM: Adaptive Differential Pulse Code Modulation) の音声圧縮方式に適用して有用な技術に関する。

【0002】

【従来の技術】 従来、音声データを圧縮するために様々な方式が開発されている。例えば、音声信号を所定周波数でサンプリング (標本化) するとともに、各時点でのサンプリングデータの値 (標本値) をその時点より前の波形から予測するとともに、この予測値と実際のサンプリングデータの値との差分をとって、この差分値を量子化および符号化することで音声圧縮を行う ADPCM と呼ばれる圧縮方式がある。また、ADPCM の圧縮方式にも、上記差分値の符号化方法により幾つもの種類がある。また、所謂 MP3 (Motion Picture Experts Group - Layer3) の音声圧縮方式などは、連続する音声信号を所定区分ごとに分割し、この分割された音声信号を周波数分解して、その結果得られたパラメータを符号化することで音声圧縮を行う。

【0003】 上記のように何れの音声圧縮方式についても、最終的に何らかのパラメータを符号化することで音声圧縮を行うので、この符号化による圧縮率を変化させることで音声データの圧縮率を変化させることが出来る。また、同一の音声圧縮方式であれば圧縮率を高くすればそれに伴って音質が悪くなる。

【0004】 ところで、音声信号には、例えば不規則な変化が多くて音声圧縮しにくい部分と、規則的な変化が多くて音声圧縮しやすい部分とが混在する。そして、音声圧縮しにくい部分において圧縮率を高くすると音質の劣化が大きくなるが、音声圧縮しやすい部分において圧縮率を高くしてもあまり音質には影響しない。このような性質を利用して、例えば MP3 などの音声圧縮方式などでは、各分割区間の音声信号の内容に応じて、音質の劣化が大きくなるような部分については圧縮率を低くする一方、音質にあまり影響しないような箇所については圧縮率を高くするようにした可変ビットレートと呼ばれる方式を採用する場合がある。このような方式によれば、音質の劣化が少なく且つ全体的なデータ圧縮率を高めることが出来る。

【0005】

【発明が解決しようとする課題】 しかしながら、上記の可変ビットレートの方式は、音声信号を複数の区間に区分けして圧縮率を変動させるため、MP3 のように元々音声信号を複数区間に区分けして扱うような方式に対しては調和性があるが、ADPCM のように音声信号を連続的に扱う圧縮方式に対しては、そのままでは調和性がなく、可変ビットレートの方式を上手く適用することが

難しい。また、MP3などの音声圧縮方式は、音声信号の周波数変換を行うなどADPCMの方式に較べて処理が複雑なため、高速な演算回路が必要になるなど、高コスト構造になるという課題がある。

【0006】この発明の目的は、比較的単純な処理で音声圧縮が行えるADPCMのような音声圧縮方式に可変ビットレートの方式を取り入れることで、音質の劣化が少なく且つ全体的なデータ圧縮率を高めることが可能で、且つ廉価なハードウェア構成でも実行することが可能な音声圧縮方法および伸長方法を提供することである。この発明の前記ならびにそのほかの目的と新規な特徴については、本明細書の記述および添付図面から明らかになるであろう。

【0007】

【課題を解決するための手段】本願において開示される発明のうち代表的なものの概要を説明すれば、下記のとおりである。すなわち、音声信号を標本化した後、各時点の標本値をそれ以前の標本値に基づき予測するとともに、この予測した値と実際の標本値との差分を量子化および符号化して音声圧縮データを生成する音声圧縮方法において、一続きの音声信号を微小区間ごとに分け、該微小区間の符号化誤差の平均値が所定の範囲に収まるように各微小区間の量子化を行うとともに、符号化誤差の平均値が所定の範囲を越えてしまう場合に量子化ビット数を変動させるように構成する。このような手段によれば、比較的符号処理が単純なADPCM方式の音声圧縮方式に、可変ビットレートの方式を適用することが可能であり、それにより音質を低下させずに全体的な圧縮率の向上が図れる。

【0008】

【発明の実施の形態】以下、本発明の好適な実施例を図面に基づいて説明する。図1は、本発明に係る音声圧縮方法および音声伸長方法を適用したボイスレコーダ10の概略構成を示すブロック図である。同図において、11は例えば単結晶シリコンのような1個の半導体チップ上に形成されたワンチップマイクロコンピュータ、12は所定の制御プログラムに従って各種演算処理を行うCPU (Central Processing Unit)、13は所定のサンプリング周波数でアナログ音声信号をデジタルに変換するAD変換器、14はデジタル音声信号をアナログに変換するDA変換器、15は音声の圧縮伸長プログラムが格納されたROM (Read Only Memory)、16はCPU12の作業メモリ空間や圧縮音声データの記憶領域を提供する内蔵RAM (Random Access Memory)、18は圧縮音声データを記憶する外部メモリ、19と20は音声入力用のマイクおよびフィルタアンプ、21と22は音声出力用のフィルタアンプおよびスピーカである。

【0009】ボイスレコーダ10は、マイク19から入力した音声信号をROM15に格納された音声圧縮プログラムを用いて圧縮し、内蔵RAM16或いは外部メモ

リ18に一旦蓄積する一方、これらのメモリに蓄積した音声圧縮データをROM15に格納された音声伸長プログラムを用いて伸長し、スピーカ22から出力したりする音声処理装置である。ROM15に格納されている音声圧縮プログラムは、ADPCM方式と可変ビットレート方式とを融合した圧縮方式を採用したものである。

【0010】図2には、この音声圧縮方式の概要を説明する図を示す。ADPCM方式は、AD変換器13で標本化されたデジタル音声信号の各時点の標本値を、それ以前の信号波形から所定の解析手法を用いて予測するとともに、その予測値と実際の標本値との差分を演算し、その差分値を指定の量子化ビット数で量子化するとともに、さらに例えばランレングスやハフマン符号化などの符号化処理を行って音声圧縮データを生成するものである。このようなADPCM方式の音声圧縮方式では、例えば図2の音声波形中、音の変化量が中程度の区間①と、音の変化量が多い区間②と、音の変化量が少ない区間③とで、所定の音質を得るために必要な量子化ビット数は異なってくる。例えば、音の変化量が多い区間②では量子化ビット数を大きくしないと音質劣化が大きくなるが、音の変化量が少ない区間③では量子化ビット数を小さくしても音質劣化はあまり生じない。

【0011】そこで、この実施例に係る音声圧縮方式では、上記差分値の量子化ビット数が音声信号に応じて自動的に変動される可変ビットレートの機能を備えたものになっている。連続的に音声信号を扱うADPCM方式の音声圧縮方式に可変ビットレートを適用する手法は次の通りである。すなわち、まず、音声信号を一定長の微小区間に分け、音声信号をADPCM方式で符号化していくとともに、音声信号を符号化する際に生じる誤差（符号化誤差）の平均を各微小区間ごとに算出し、この符号化誤差の平均値が所定範囲から逸脱した場合に、上記量子化ビット数を変動させて、全ての微小区間において符号化誤差の平均値が所定範囲内に収まるようにしたものである。

【0012】次に、この音声圧縮方式の詳細をフローチャートに基づき説明する。図3には、ボイスレコーダ10において実行される音声圧縮処理のフローチャートを示す。この音声圧縮処理が開始されると、まず、ステップS1において、音声圧縮データのヘッダ情報に例えば音声信号の開始の振幅点など必要なデータを設定する。次に、ステップS2において音声圧縮処理を行う音声信号の微小区間を設定する。

【0013】ステップS3、S4では、従来のADPCMの圧縮処理と同様に、音声信号中の処理対象の標本値（サンプリングデータ）を前の信号波形から予測するとともに、この予測値と実際の標本値との差分を指定の量子化ビット数で量子化し、さらに所定の符号化処理で符号化して、圧縮音声データの要素データを生成する。ステップS5では、符号化誤差として、圧縮音声データが

ら復調した差分値と実際の差分値との差の絶対値を演算し、先の要素データとこの符号化誤差のデータとをRAM16等に保存する。

【0014】そして、ステップS6において、先のステップS2で設定された微小区間の終了点か否か判別し、終了点でなければステップS3に戻り、終了点であればステップS7に移行する。それにより、上記ステップS3～S5の処理が微小区間の開始点から終了点まで繰り返し行われ、1個の微小区間について完了したらステップS7に移行する。ステップS7では、ステップS5で保存されてきた符号化誤差のデータに基づき、当該微小区間における符号化誤差の平均値を演算する。そして、続くステップS8において、この符号化誤差の平均値から最適量子化ビット数を決定し、次の微小区間の量子化に用いるビット数とする。

【0015】ここで、最適量子化ビット数とは、符号化誤差の平均値が所定範囲内に収まるように設定されたビット数のことである。この範囲の上限值および下限値の設定は、量子化に用いる量子化ビット数の種類や音質等を考慮して適宜行えばよいが、量子化ビット数ごとに異なる値としても良いし、全量子化ビット数で共通の値としても良い。例えば、最適量子化ビット数とされる範囲として、量子化ビット数が4ビットの場合には符号化誤差の平均値 X_4 が $X_4 \geq 5$ 、量子化ビット数が3ビットの場合には符号化誤差の平均値 X_3 が $8 \geq X_3 \geq 3$ 、量子化ビット数が2ビットの場合には符号化誤差の平均値 X_2 は $5 \geq X_2 \geq 0$ などと設定することが出来る。

【0016】そして、ステップS8において、符号化誤差の平均値が所定範囲に収まっていれば、現在の量子化ビット数を最適なものとし、所定範囲に収まってなく、平均値が大きい方に外れていればそれに応じて量子化ビット数を大きくする一方、平均値が小さい方に外れていればそれに応じて量子化ビット数を小さくして、最適な量子化ビット数とする。ステップS9では、量子化ビット数が変更されたか判別し、変更されなければステップS2に戻り、変更されれば次のステップS10に移行する。それにより、最適な量子化ビット数がずっと同じであれば、複数の微小区間においてステップS2～ステップS8が繰り返される。

【0017】一方、量子化ビット数が変更されてステップS10に移行した場合には、量子化ビット数の変更が生じた微小区間を除き、それ以前の微小区間で同一の量子化ビット数で量子化されてきた区間の長さ、その量子化ビット数とをRAM16等に保存して音声圧縮データのヘッダ情報に含める(ステップS10)。そして、この区間の音声圧縮データを確定データとしてRAM16等に保存する(ステップS11)。

【0018】ステップS12では、音声データが終了か否か判別して、まだ終りでなければステップS1に戻って、ステップS8で最適量子化ビット数が変更された微

小区間から、変更後の量子化ビット数を用いて、同様の処理を繰り返す。一方、音声データの終りであれば、この音声圧縮処理を終了する。以上のように、この音声圧縮処理プログラムによれば、ADPCM方式と可変ビットレート方式とが融合した音声圧縮方式により、音質の劣化が少なく、且つ、データ量も小さい音声圧縮データを生成することが出来る。

【0019】図4には、上記ボイスレコーダ10により実行される音声伸長処理のフローチャートを示す。この音声伸長処理は、上記音声圧縮処理で生成された音声圧縮データを伸長するものである。この音声伸長処理が開始されると、まず、ステップS21において音声圧縮データのヘッダ情報から、同一の量子化ビット数で量子化されている区間長の情報とその量子化ビット数の情報とを読み出す。

【0020】次いで、音声圧縮データから、時系列に配置されている要素データを読み出し(ステップS21)、この要素データを復調するとともに、前の波形に基づき予測される点に復調した差分値を加算して音声データに伸長する(ステップS23)。そして、この音声データを一旦RAM16等に保存し(ステップS24)、その後、DA変換器14を介してスピーカ22から出力される。その後、ステップS25において、ステップS21で読み出した区間の終了に達したか否か判定し、まだ、終了に達していなければステップS22に戻り、終了であれば次のステップS26に移行する。それにより、同一の量子化ビット数で量子化されている区間においてステップS22～S24の伸長処理が繰り返し実行され、該区間の音声データが伸長される。

【0021】ステップS26では、音声圧縮データの終了に達したか否かを判定し、終了に達していなければステップS21に戻って次の区間の伸長処理を開始し、終了に達したならばこの伸長処理を終了する。以上のように、この実施例の音声圧縮方法および音声伸縮方法によれば、基本的な圧縮方式はADPCMなので、音声信号を周波数分解する方式に較べて処理が単純でありハード構成のコスト低減を図ることが出来る。また、可変ビットレートの方式を取り入れたことで、音質を余り劣化させずに全体的な圧縮率を高くすることが出来る。また、音声信号を圧縮する際の遅延は、最大でおおよそ微小区間の信号入力にかかる時間となるので、それほど大きくならない。また、伸長時の遅延は伸長処理にかかる動作時間のみのため、非常に小さくすることが出来る。

【0022】以上本発明者によってなされた発明を実施例に基づき具体的に説明したが、本発明は上記実施例に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。例えば、実施例では、差分値の符号化誤差は、音声圧縮データから復調した差分値と実際の差分値との差の絶対値として説明したが、例えば、差の2乗値とするなど、その表わし

方は適宜変更可能である。また、量子化で生じる誤差と符号化で生じる誤差とを合算したものを符号化誤差としても良いし、量子化で生じる誤差のみを符号化誤差として量子化ビット数の変更の判断に用いるようにしても良い。

【0023】また、同一の量子化ビット数で連続的に量子化が行われている区間を示す区間長情報とその量子化ビット数を示す情報の格納位置は、全区間の情報をまとめて音声圧縮データの先頭に配置しても良いし、或いは、各区間の音声圧縮データの先頭にそれぞれ配置する

ようにしても良い。

【0024】以上の説明では主として本発明者によってなされた発明をその背景となった利用分野であるボイスレコーダへ適用した例について説明したがこの発明はそれに限定されるものでなく、次に示すような音声処理装置に広く利用することが出来る。

【0025】図5には、本発明の第2の適用例である音声再生装置の概略構成図を示す。図1のボイスレコーダ10がマイク19から音声信号を入力していたのに対して、この音声再生装置30は外部メモリ31等から音声データを

入力し、この音声データを音声圧縮して内蔵RAM16等に格納するように構成されたものである。このような音声再生装置30においても、外部から入力した音声データを圧縮する際やこの圧縮データを伸長する際に本発明に係る圧縮方法および伸長方法を適用することが出来る。

【0026】図6には、本発明の第3の適用例である音声合成装置の概略構成図を示す。この音声合成装置40は、音声合成により音声出力を行う装置であるが、音声合成に使う音源データとして、外部メモリ41から圧縮して内蔵RAM16に格納された圧縮音源データを使用するものである。音声合成の際には、内蔵RAM16の圧縮音源データの中から必要な圧縮音源データが伸縮されて音声合成制御部42に供給され、該音声合成制御部42においてこれらの音源データが合成されてスピーカ22から出力される。このような音声合成装置40においても、外部メモリ41の音源データの圧縮並びに内蔵RAM16の圧縮音源データの伸長に本発明に係る圧縮

方法および伸長方法を適用することが出来る。

【0027】

【発明の効果】本願において開示される発明のうち代表的なものによって得られる効果を簡単に説明すれば下記のとおりである。すなわち、本発明に従うと、音声信号の標本値を前の波形から予測しその差分を量子化および符号化する音声圧縮方式なので、例えば音声信号を周波数分解して圧縮する方式に較べて処理が単純でありハード構成のコスト低減を図ることが出来る。さらに、差分値の量子化ビット数を音声信号の変化の割合に応じて可変にしたことで、音声信号の音質を余り劣化させずに全体的な圧縮率を高くすることが出来るという効果がある。

【図面の簡単な説明】

【図1】本発明の実施例である音声圧縮方法および音声伸長方法を適用したボイスレコーダの概略構成を示すブロック図である。

【図2】実施例の音声圧縮方法の概要を示す説明図である。

【図3】図1のボイスレコーダにおいて実行される音声圧縮処理の手順を示すフローチャートである。

【図4】図1のボイスレコーダにおいて実行される音声伸長処理の手順を示すフローチャートである。

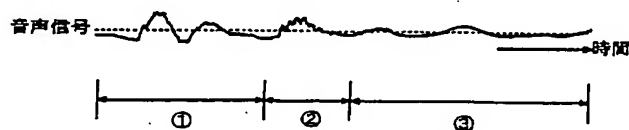
【図5】本発明の第2の適用例である音声再生装置の概略構成を示すブロック図である。

【図6】本発明の第3の適用例である音声合成装置の概略構成を示すブロック図である。

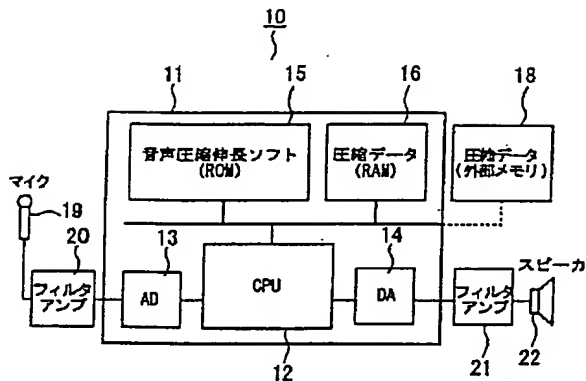
【符号の説明】

10	ボイスレコーダ
12	CPU
13	AD変換器
14	DA変換器
15	ROM
16	内蔵RAM
18	外部メモリ
19	マイク
22	スピーカ

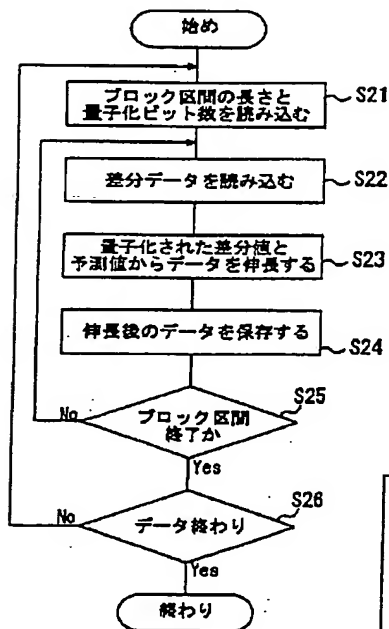
【図2】



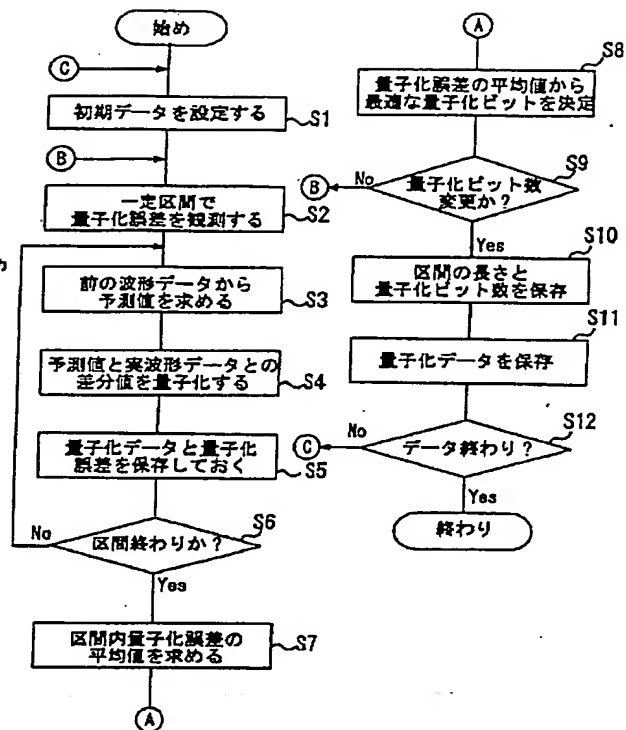
【図1】



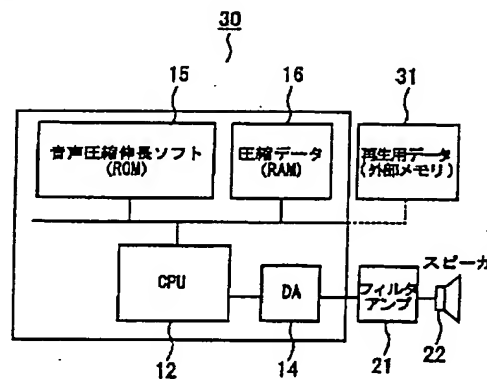
【図4】



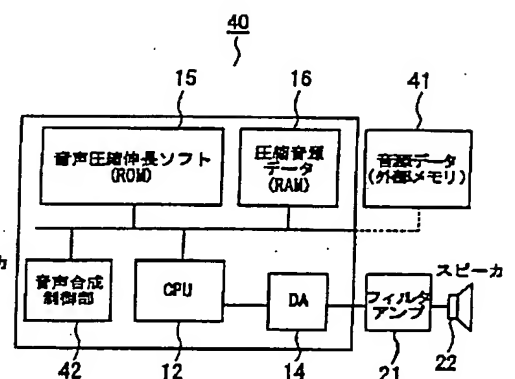
【図3】



【図5】



【図6】



フロントページの続き

(72)発明者 瀬尾 充彦
東京都小平市上水本町5丁目22番1号 株式会社日立超エル・エス・アイ・システムズ内

(72)発明者 田中 誠
東京都小平市上水本町5丁目22番1号 株式会社日立超エル・エス・アイ・システムズ内

F ターム(参考) 5D045 DA07
5J064 AA02 BA05 BB03 BC01 BC06
BC07 BC15 BC16 BC26 BC29
BD00